Transactions of the VŠB – Technical University of Ostrava, Mechanical Series
No. 1, 2009, vol. LV
article No. 1675

Petr ŠPAČEK*, Petr NOVÁK, Vladimír MOSTÝN*****

PROVIDING THE 3D PERCEPTION TO OPERATOR

ZPROSTŘEDKOVÁNÍ 3D VJEMU OPERÁTOROVI

Abstract

The article describes the realised system of stereovision providing space perception to operator. Final appliance is focused to using on the service mobile rescue robot. The operator has 3D helmet putted on and the scenery is taken by pair of parallel cameras. The software transformation of both pictures is implemented for every eye. Article introduced cameras calibration, camera distortion removing, automatic finding distance of focusing, transformation and filtering.

Abstrakt

Příspěvek popisuje realizovaný stereosystém, který zprostředkovává prostorový vjem operátorovi. Výsledné zařízení je určeno k použití na servisním mobilním zásahovém robotu. Operátor má nasazeny 3D brýle a scéna je snímaná dvěma rovnoběžným kamerami. Za použití zhotoveného software, který je napsán v jazyce C++, je provedena transformace obrazu tak aby každému oku příslušel vhodný obraz. Článek představuje část od kalibrace kamer, přes odstranění zkreslení kamer, automatické hledání optimální vzdálenosti pro zostření systému, způsob transformace a nakonec aplikovaný filtr.

INTRODUCTION

Regarding to typical missions of service mobile robots for rescue purposes and especially in danger environment is a motion control by operator very important. The operator controls mobile robots on the basis of information obtained by sensory subsystem. The typical sensors – for these purposes – are cameras. Camera gives image information about robot's surrounding. This information is for the operated robot this greatly important. There is but one problem – the image information giving only one camera does not give three dimensional feeling. The absence of this perception can make difficult or forbid robot control. Bellow described system, based on stereo cameras, gives 3D image information.

SYSTEM DESCRIPTION

The described system is pointed for providing of the stereovision feeling to operator. Taking scene there is assured by pair of the parallel cameras DFK31F03 without possibility of cross tilt. The projection of the final images is by help of the 3D helmet of virtual reality CyberMind HI-Res900. The design of this system is in the **Figure 1**.

The computing algorithm including a necessary software is has been written under C++ language and uses the DirectX interface for working with graphics card. This configuration allowed achieving high computing speeds and building the system for real running. Part of algorithms uses free OpenCV library.

* Ing. Petr Špaček, Department of Robotics, VŠB-TU, 17. listopadu 15/2172, 70833 Ostrava-Poruba, Czech Republic

** Doc. Dr. Ing. Petr Novák, Department of Robotics, VŠB-TU, 17. listopadu 15/2172, 70833 Ostrava-Poruba, Czech Republic, ☎+420596993595, petr.novak@vsb.cz

*** prof. Dr. Ing. Vladimír Mostýn, Department of Robotics, VŠB-TU, 17. listopadu 15/2172, 70833 Ostrava-Poruba, Czech Republic

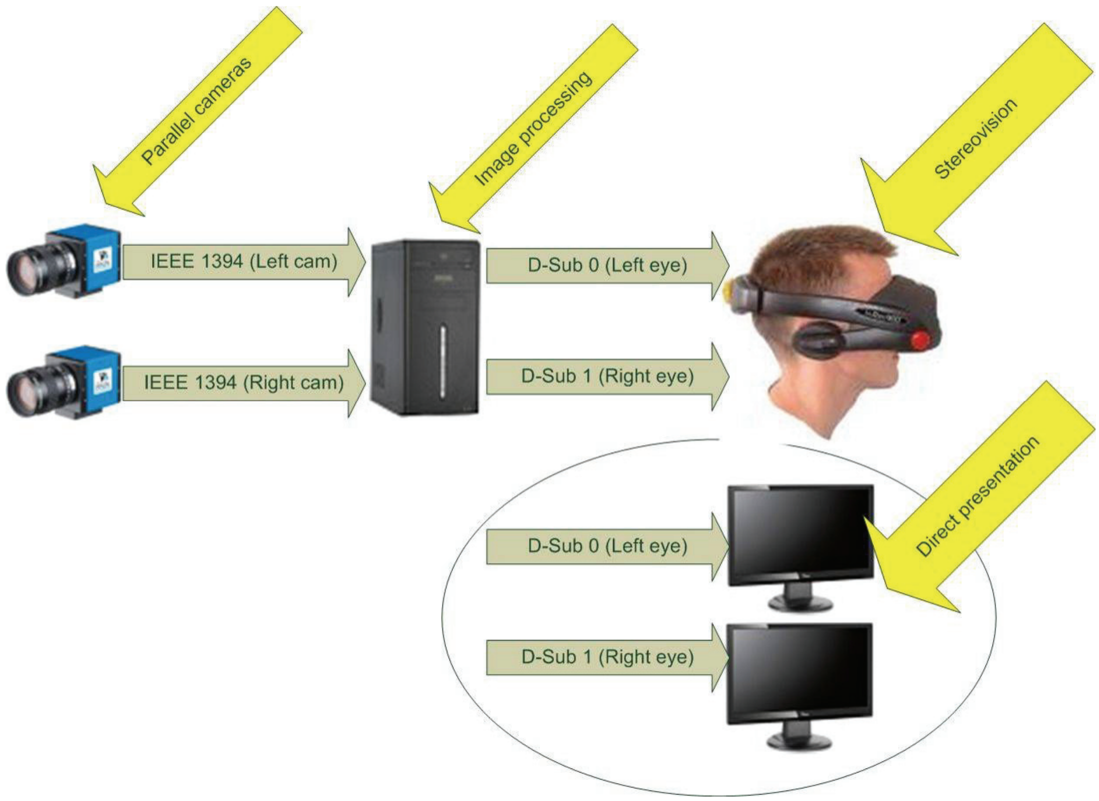

Figure 1 Design of the stereovision system

PROJECTIVE GEOMETRY OF THE CAMERA

The position relationship between point Q, with coordinates the real scene and displayed point q which has coordinates of image is called projective transformation. Next, we use homogonous co-ordinates and so-called matrix of inner camera's parameters. Equations bellows describe projection point Q of real word to point q in camera.

$$q = MQ$$

where,

$$q = \begin{bmatrix} x \\ y \\ w \end{bmatrix}, \qquad M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \qquad Q = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Subsequently f_x is focal lenght in x axis , f_y focal lenght in y axis, c_x and c_y are coordinates of the main camera point (point of optical axis camera and plane of picture crossing).

RADIAL LENS DISTORTION

The radial distortion of lens is zero in optical centre and in direction of margins increases. In praxis there is possible function of radial distortion approach by some first members of the Taylor expansion. For this members here is k_1 for first and k_2 for second used. Third (and last) member k_3 is using for lens with high distortion – as "fish eye". The point on the picture will be placed on the new position by next equations:

$$x_{nov} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$
$$y_{nov} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

where,

$$r^2 = x^2 + y^2$$

x_{nov} and y_{nov} are new point coordinates and x and y are coordinates of the original point.

TANGENTIAL LENS DISTORTION

The tangential distortion grows up by mounting and assembling of camera. This distortion is characterised another two parameters p_1 and p_2. Their meaning is evident from equations below:

$$x_{nov} = x + \left(2p_1 y + p_2(r^2 + 2x^2)\right)$$

$$y_{nov} = y + (p_1(r^2 + 2y^2) + 2p_2 x)$$

CAMERA PROPERTIES

The ascertained camera properties are presented in next tables and graphs.

Table 1 The parameters of the left camera

Parameter	**Value**
Focal length in axis x Fx	874 pixel
Focal length in axis y Fy	871 pixel
Point of optical axis crossing (coordinate X)	597 pixel
Point of optical axis crossing (coordinate Y)	340 pixel
Radial lens distortion ratio k1	-0,2243
Radial lens distortion ratio k2	0,14425
Radial lens distortion ratio k3	0,00144
Tangential lens distortion ratio p1	-0,00316
Tangential lens distortion ratio p2	0

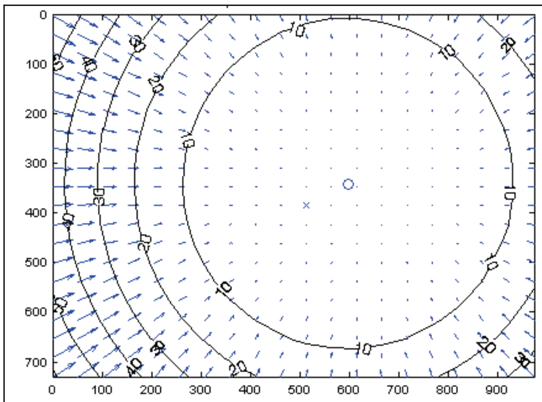

Figure 2 Left camera radial lens distortion

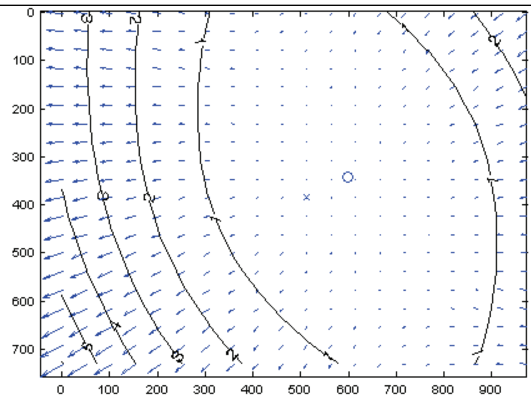

Figure 3 Left camera tangential lens distortion

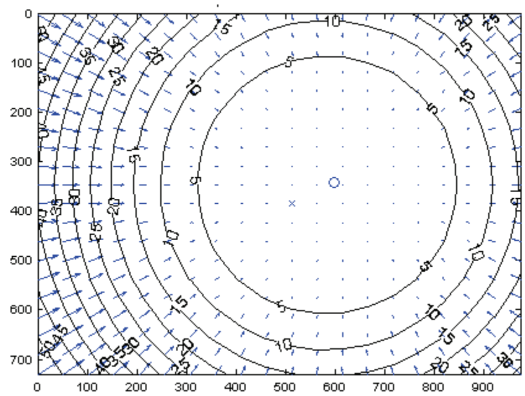

Figure 4 Complete model of the left camera lens distortion

The same way we can find the parameters for right camera.

RECTIFICATION

In real sterevision system are not images left and right cameras matched each other. This means that image sensor chips do not lie in the same plane.

Due to we must perform rectification. The aim is achieving state when images planes booth cameras are identical including positioning of the identical rows.

In the next figures the effect of rectification is presented. It is suitable notice removing the tangential and radial distortion. There is evident alignment of the rows too.

Figure 5 Before rectification

Figure 6 After rectification and rows alignment

IMAGE TRANSFORMATION DEPENDING ON FOCUSING LENGTH

The fundamental of human stereovision is based on eyes tilting against each other. The same principle can by use for camera stereovision system. But there is one problem – it is requirement on cameras mechanical tilting and its control. The mechanical part must by precise and extend the system complexity including its cost and must consist of the positioning control system for it too. Next described system does not use the mechanical motion of cameras.

The designed system stereovision is based on shifting both images against each other. In the **Figure 7** is presented the example taking images which are focused at infinite. Cameras resolution there is 1024x768 pixels and next the slices 800x600 are done (this resolution is native of 3D helmet).The centre of image cutting is identical with camera optical axes. Those cuttings are screened on displays of 3D helmet.

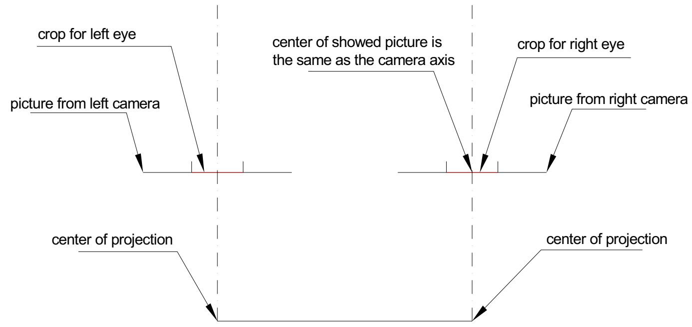

Figure 7 Focusing on infinite

In case of near object position, the system simulates cameras tilting. In the **Figure 8** is presented behaviour on “focusing” on near object. The images are taking in 1024x768 resolutions too. But both slices were cut in different places. The centre of this cutting can be calculates under equations below (the centre of y axes lies in the half native resolution).

$$S_{l_v} = \frac{x_l - x_p}{2} + S_l$$

$$S_{p_v} = S_p - \frac{x_l - x_p}{2}$$

Where S_{l_v} and S_{p_v} are centres of left and right slices respectively, x_l and x_p are image coordinates of observed point, S_l and S_p are image coordinates the centre of image taking by camera. The difference $x_l - x_p$ is called disparity. In case of the identical resolution both cameras, we can write the equation by matrix form:

$$\begin{bmatrix} S_{l_v} \\ S_{p_v} \end{bmatrix} = \begin{bmatrix} 0.5 & 1 \\ -0.5 & 0.5 \end{bmatrix} \begin{bmatrix} d \\ R \end{bmatrix}$$

The d is disparity and R is camera resolution.

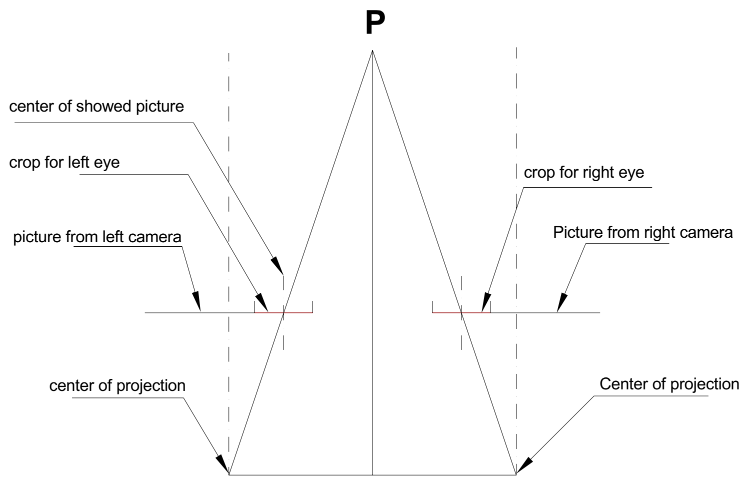

Figure 8 Focusing on the near objects

FINDING THE MATCHING POINTS

The used method is called SAD - Sum of Absolute Differences. In surroundings of point in left image, quarter little window is defined – in our case the size is 21x21 pixels. Its centre lies in point which coordinate (x coordinate we know, we find only x coordinate) we need find in second image. In the **Figure 9** is this little window illustrated. The calculation is based on this equation below:

$$diference(x_p, y) = \sum_{\Delta x=-a}^{a} \sum_{\Delta y=-a}^{a} abs[J_1(x_l + \Delta x, y + \Delta y) - J_2(x_p + \Delta x, y + \Delta y)]$$

Where J_1 and J_2 are brightness pixel with coordinates (x, y) for left and right images,

a is half of looking up window.

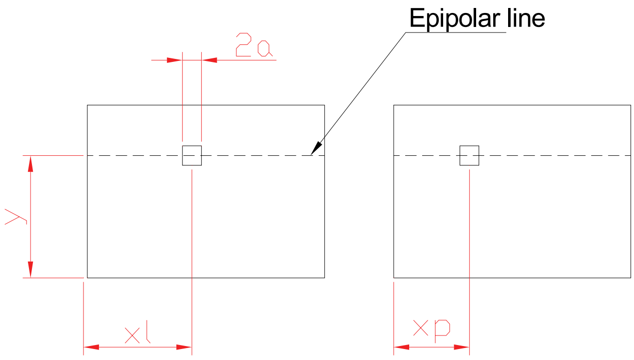

Figure 9 SAD algorithm

OPTIMAL DISPARITY DETERMINATION

The mentioned algorithm above is successful well for objects with higher surface texture. In outdoor environment are textured objects enough – opposite to indoor, where in rooms, vestibules etc. are mainly big surfaces (for example white walls) they have texture very low. Consequently, that majority of points are very similar and it is difficult recognize them. The algorithm must decide which object in the image is operator's object of interest.

The algorithm solving it is following:

1) In centre of left image is a slice 160x160 pixel – see Figure 10. In area of interest (green window) the points found by algorithm are marked gray colour. The black colour matches points not found in the right image.
2) Now we can determine the optimal disparity by help of disparity histogram.
3) In histogram the maximal value is founded which is responding to biggest occurrence and it is selected as optimal disparity.

Figure 10 Region searching (in the left Image)

The **Figure 11** presents histogram of disparities for near object – see **Figure 13**. Axes x represents the value of disparity and axes y occurrence identical disparity.

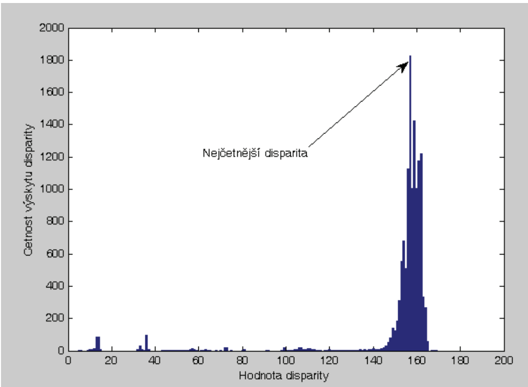

Figure 11 Histogram of disparities for near object

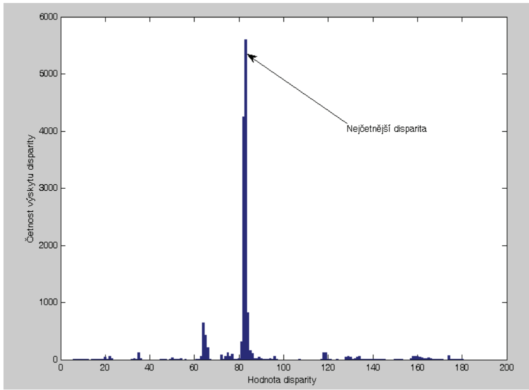

Figure 12 Histogram of disparities for far object

Figure 13 Images of near object

Histogram of disparities for far object is showed in the **Figure 12**. There is obvious that most frequently disparities are lower. It corresponds to larger distance to object. The corresponding images of this scene you can see in the Figure 14.

In real situations can come situation, when maximal disparity in histogram more bars consist. In the case can be chosen every time other bar and sterevision images will be oscillating. Due to there is special “disparity” filter used.

Figure 14 Images of far object

CONCLUSION

The presented article described solving of stereovision subsystem of mobile robot with immovable cameras. The stereovision perception is achieved by sifting of images slices and has system of automatic focusing on the point of interest.

The system was tested in various situations ant results confirmed its applicability in real situations. Next tasks there are software optimizing, his porting into small embedded computer including addition graphical and text information about other robot’s subsystems as temperature, its tilt, power status, image of rear camera etc.

This article was compiled as part of projects FT-TA3/014, supported by the Fund for University Development from the Ministry of Industry and Trade.

REFERENCES

[1] Mostýn V. *Mechatronika* - 1. vyd.. - Ostrava : VŠB - Technická univerzita Ostrava, 2000 - 68 s. ISBN 80-7078-734-1

[2] Sojka E., *Digitální zpracování a analýza obrazů* - 1. vyd. - Ostrava : VŠB - Technická univerzita Ostrava, 2000 - 133 s. ISBN 80-7078-746-5

[3] Z. Zhang, *Flexible camera calibration by viewing a plane from unknown orientations.,* Microsoft research technical report 1998, http://research.microsoft.com/˜zhang

[4] D. C. Brown*, Close-range camera calibration,* 1971

[5] Heikilla J., Silvén, O. *A four-step camera calibration procedure with implicit image correction*, 1997

[6] Tsai R. Y., *A versatile camera calibration technique for high accuracy 3D machine vision metrology using off -the-shelf TV cameras and lenses*, 1987

[7] Bouguet, J.Y., *Camera calibration toolbox*, dostupný z www: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html,

[8] *OpenCV Wiky, Open Source Computer Vision Library Wiki,* http://opencvlibrary.sourceforge.net

[9] Tvarůžka, A. *Senzorický subsystém robotu.* PhD thesis. Department of robototics, VŠB-TUO, Ostrava, 139 s., 2008.